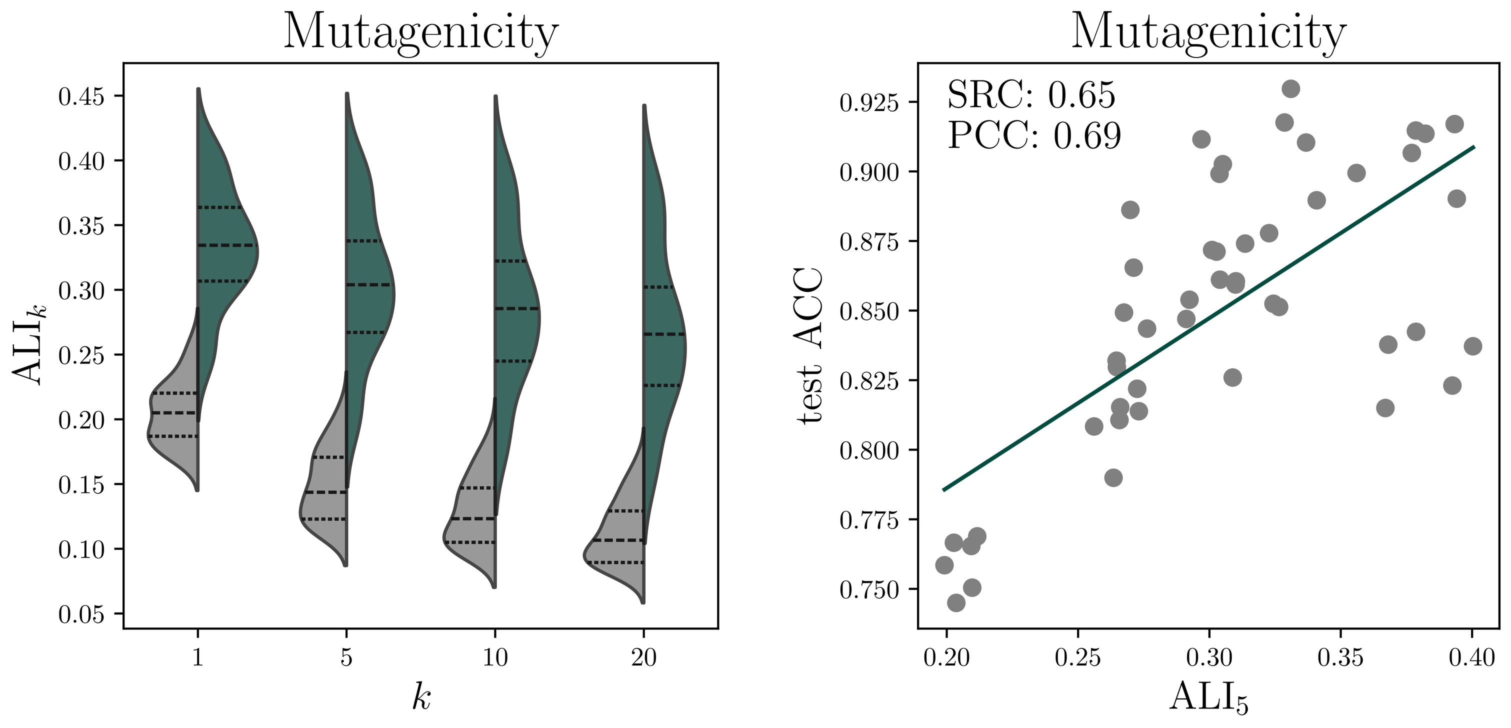
WILTing Trees: Interpreting the Distance Between MPNN Embeddings

Masahiro Negishi, Thomas Gärtner, Pascal Welke

MPNN Embedding Distance Aligns With Target Label Distance



$$\mathrm{ALI}_k(d_{\mathrm{MPNN}}, d_{\mathrm{func}}) := \frac{1}{|\mathcal{D}|}\sum_{G\in\mathcal{D}}\left[-\frac{\sum_{H\in\mathcal{N}_k(G)} d_{\mathrm{func}}(G,H)}{k} + \frac{\sum_{H\in\mathcal{D}\setminus(\mathcal{N}_k(G)\cup\{G\})} d_{\mathrm{func}}(G,H)}{|\mathcal{D}|-k-1}\right]$$

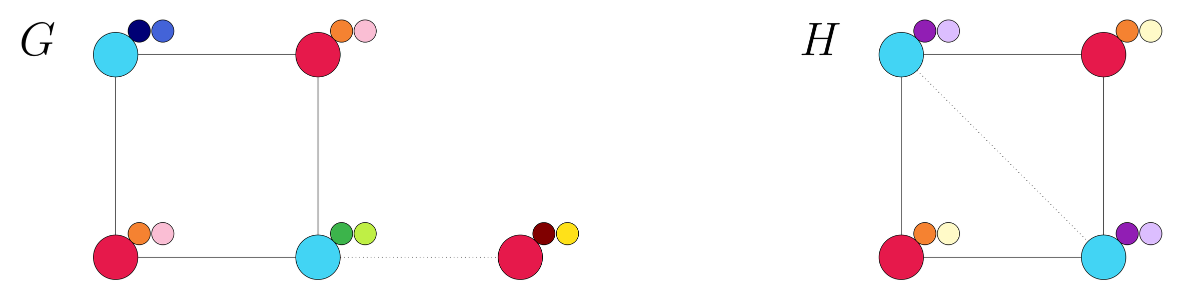
The MPNN embedding distance (d_{MPNN}) aligns with the distance between target labels (d_{func}) after training with CE or RMSE loss.

The stronger the alignment, the higher the performance.

Q: How do MPNNs embed graphs in a way that respects the distance between target labels?

MPNN Embedding Distance Is a Distance Between Weisfeiler Leman (WL) Color Multisets

Weisfeiler Leman algorithm:



E.g.) ● = HASH(●, {{(●, —), (●, —)}}) and ● = HASH(●, {{(●, —), (●, —), (●, —)}}).

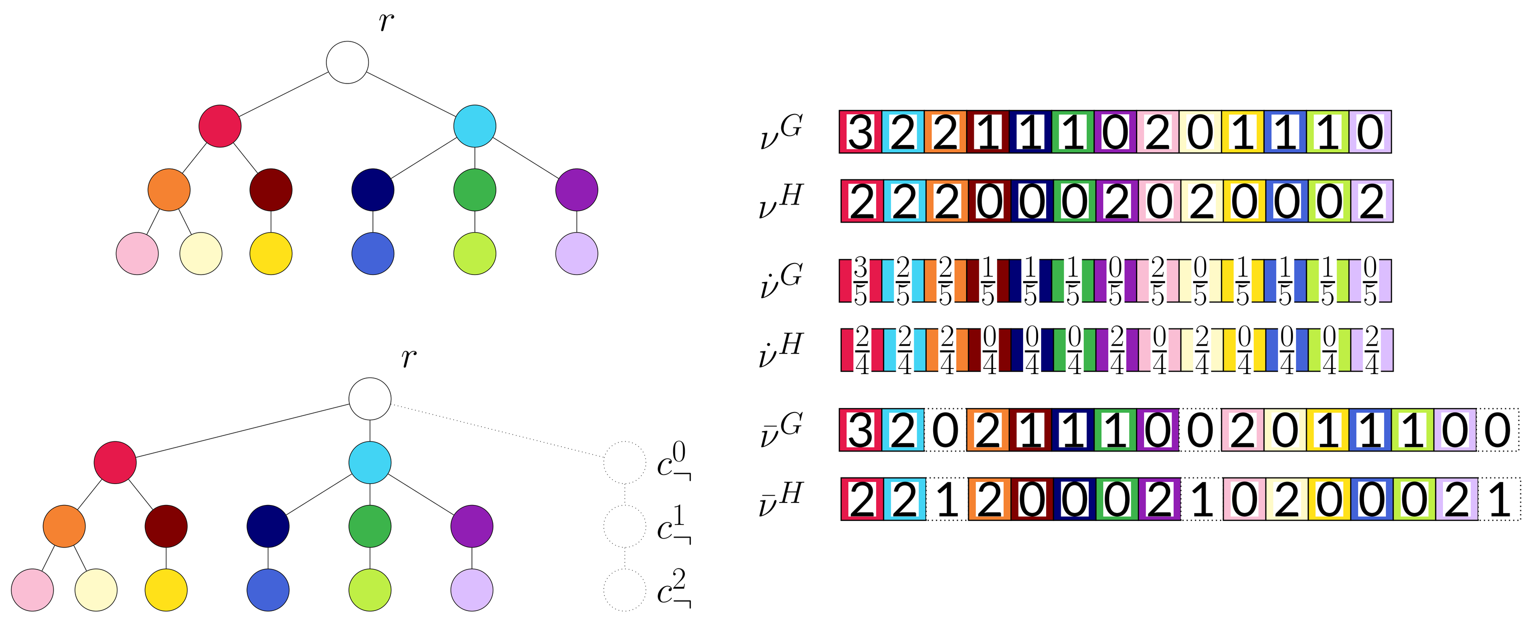
The expressive power of the WL algorithm bounds that of MPNNs.

Thus, there exists a function f s.t.

$$\begin{aligned} d_{\mathrm{MPNN}}(G,H) &:= \|\mathrm{MPNN}(G) - \mathrm{MPNN}(H)\|_2 \\ &= \|f(\{\!\{c_v^{(L)} \mid v \in V_G\}\!\}) - f(\{\!\{c_v^{(L)} \mid v \in V_H\}\!\})\|_2 \\ &=: d_f(\{\!\{c_v^{(L)} \mid v \in V_G\}\!\}, \{\!\{c_v^{(L)} \mid v \in V_H\}\!\}). \end{aligned}$$

Approximate d_{MPNN} with a more interpretable distance between WL color multisets!

Weisfeiler Leman Labeling Tree (WILT) Distance Between WL Color Multisets Is Interpretable and Efficiently Computable



Two equivalent definitions of the WILT distance:

$$\begin{cases} d_{\mathrm{WILT}}(G,H;w) := \min_{P\in\Gamma} \sum_{v_i\in V_G}\sum_{u_j\in V_H} P_{i,j}\, d_{\mathrm{path}}(c_{v_i}^{(L)}, c_{u_j}^{(L)}) & (\Gamma: \text{Set of valid transports}) \\ d_{\mathrm{WILT}}(G,H;w) := \sum_{c\in V(T_{\mathcal{D}})\setminus\{r\}} w\left(e_{\{c,p(c)\}}\right) \left|\nu_c^G - \nu_c^H\right| \end{cases}$$

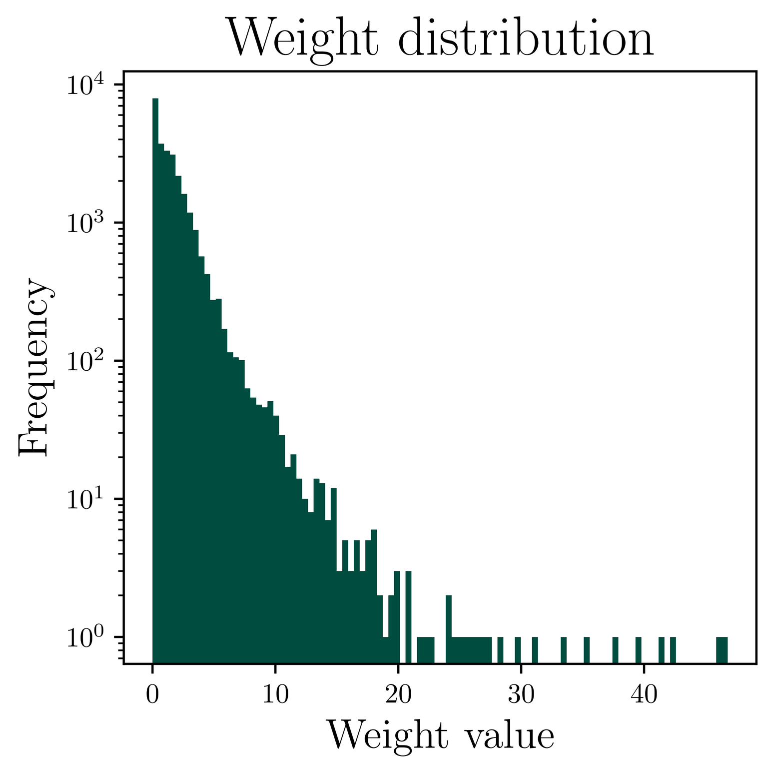
d_{WILT} can approximate d_{MPNN} by adjusting edge weights w (distillation).

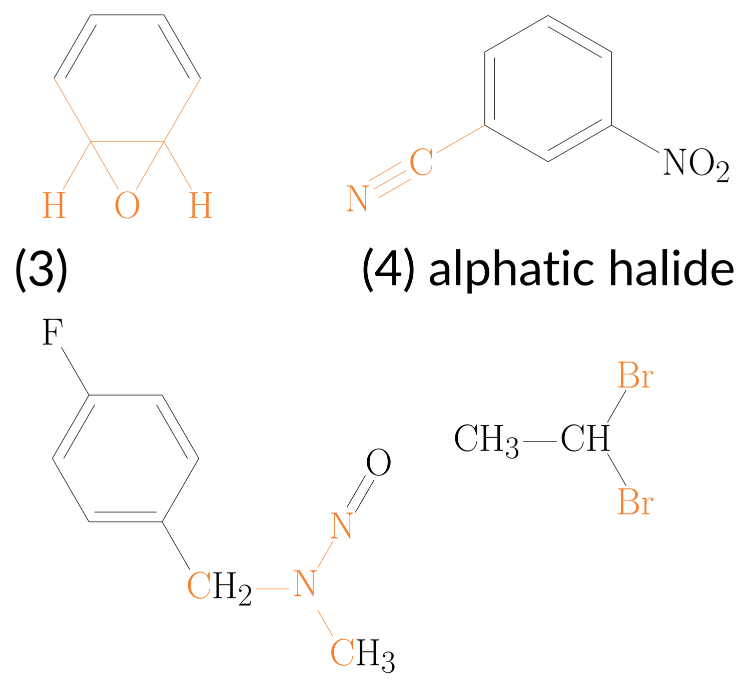
WL colors c with large weights mostly determine d_{WILT}, thus d_{MPNN}.

$d_{\mathrm{WILT}}(G,H;w)$ can be computed in $O(|V_G| + |V_H|)$.

The WILT distance generalizes the distances corresponding to well-known graph kernels. [KGW16; Tog+19]

MPNN Embedding Distance Is Determined By Only Small Number Of WL Colors

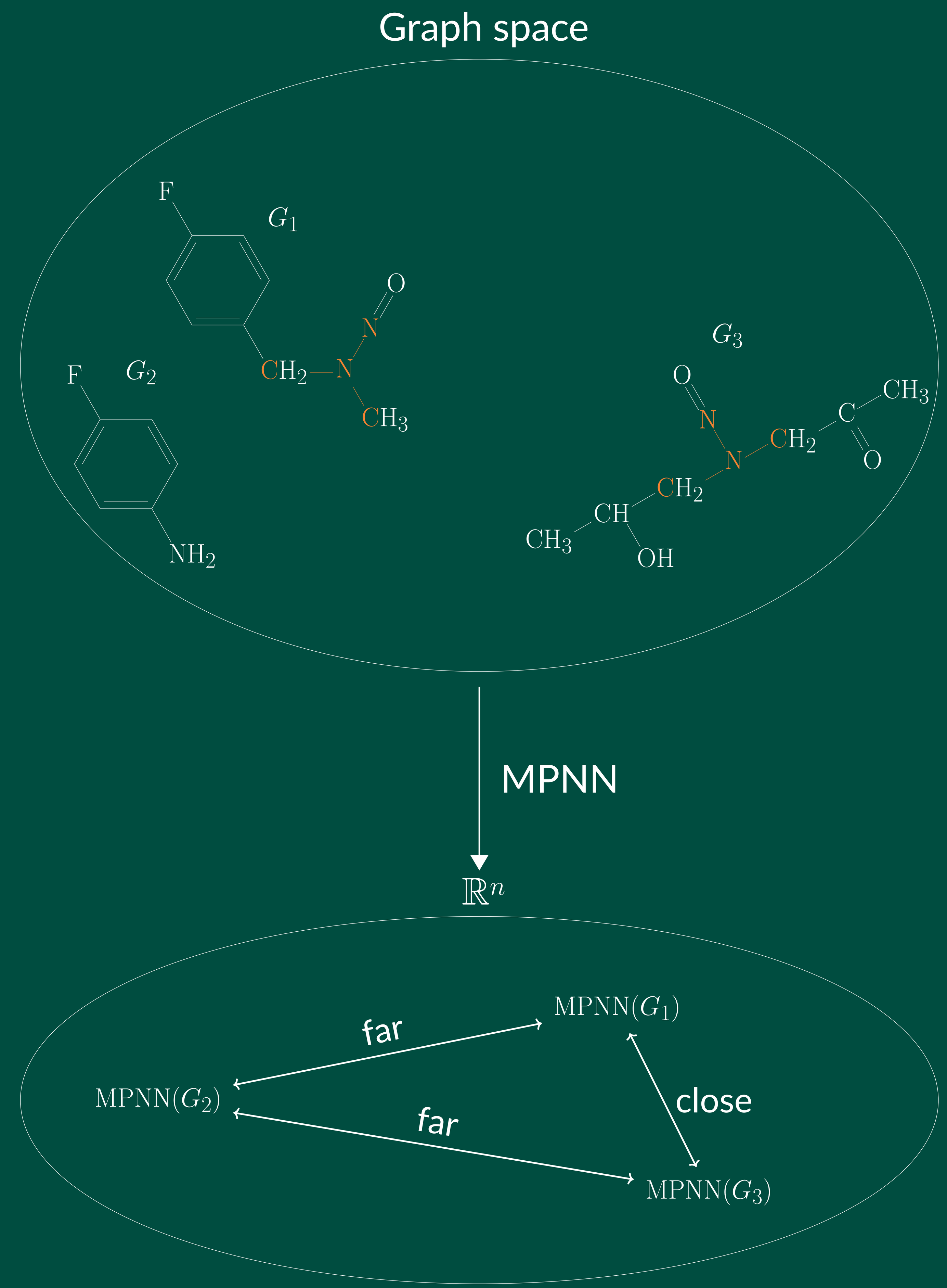




GCN trained on the Mutagenicity dataset, which contains 2401 mutagens and 1936 non-mutagens.

Identified colors agree with chemical studies. [KMB05]

The relative position of MPNN embeddings is determined by a **small** set of subgraphs.



References

[KMB05] Jeroen Kazius, Ross McGuire, and Roberta Bursi. "Derivation and validation of toxicophores for mutagenicity prediction". In: *Journal of Medicinal Chemistry* 48.1 (2005), pp. 312–320.

[KGW16] Nils M. Kriege, Pierre-Louis Giscard, and Richard C. Wilson. "On Valid Optimal Assignment Kernels and Applications to Graph Classification". In: *Advances in Neural Information Processing Systems*. 2016.

[Tog+19] Matteo Togninalli et al. "Wasserstein Weisfeiler-Lehman graph kernels". In: *Advances in Neural Information Processing Systems* (2019).